Emily Hough-Kovacs

+1 (603) 484-2534 | San Francisco, CA | emilyhoughkovacs@gmail.com
linkedin.com/in/emilyhoughkovacs | www.emilyhk.com | github.com/emilyhoughkovacs

Summary

Senior data scientist with 10+ years at Reddit, Meta, and Spotify, now transitioning full-time into AI safety. Track record of building from zero: Spotify's creator-side metrics suite, adopted as source of truth across teams, and the measurement foundation for Reddit's ads attribution domain, which had none. Turns ambiguous problems into measurable ones, partners across product, engineering, and business, and has hosted AI safety community events in San Francisco.

AI Safety

Mercial Research Fellowship — writing with three co-authors on the current state of AI safety. Oct – Nov 2026
BlueDot Technical AI Safety Course — six-day intensive introduction to technical AI safety & alignment. Oct 2026
Nightingale — work trial: triaged follow-up report from the collusion.wiki agent swarm investigation. Sep – Oct 2026
Community Building — co-hosted an AI Research Circle session on Constitutional AI. Mar 2026
hosted a screening and discussion of *The Thinking Game*. Jan 2026

Projects

research-watcher — daily LLM pipeline (Python, Claude API, GitHub Actions) that ranks and summarizes new AI papers.
Agentic Customer Behavior Simulator — persona-based agents model real customer purchase intent and behavior.
Build Your Own LLM — writeup from a hands-on workshop building a GPT-style model from scratch (PyTorch).

Experience

Reddit — Senior Data Scientist, Ads Marketplace March 2026 – October 2026

- Owned data science for the identity matching and attribution domain, standing up its measurement foundation (metric taxonomy, pipelines, reporting) in a system that had no true measurement; validated engineering's opportunity-sizing against source data.
- Built two Airflow ETL pipelines powering an internal dashboard for ads-engineering data amendments and make-goods during account incident remediation, keeping billed and reported figures consistent while cutting manual ticket overhead.
- Built a supply-and-demand monitoring dashboard surfacing contextual ad-to-post relevance signals across the marketplace, informing inventory and pacing decisions.

Meta — Senior Data Scientist, Reality Labs July 2024 – August 2025

- Built and evaluated a random forest ensemble predicting user purchase intent, assessing performance tradeoffs.
- Built long-range forecasts for Lifestyle and Entertainment teams' key metrics, informing roadmap and resourcing decisions.
- Ran observational analyses of MR vs. fully immersive VR effectiveness, quantifying the impact of physical space on user engagement.

Spotify — Data Scientist Aug 2022 – March 2024
May 2016 – May 2021

- Pioneered the metrics suite for the creator side of the music marketplace, adopted as the source of truth across creator-facing teams.
- Built monitoring for time-to-delivery of podcast content, surfacing supply-chain bottlenecks for podcast partnerships.
- Produced creator-side insights for Spotify Wrapped '22 and '23 (top editorial tracks, artists, albums), surfaced in Spotify's public, consumer-facing product.

Tumblr — Product Analyst September 2014 – October 2015

- Extracted business insights for a platform serving 250M+ blogs and 60M+ posts/day; automated reporting and built a Flask app to run and email ad-hoc Hive jobs.

Education

Barnard College, Columbia University — B.A., Mathematics / Computer Science May 2014

Skills & Tools

Languages & Libraries: Python (pandas, NumPy, scikit-learn, statsmodels, matplotlib), SQL (expert)
Statistics & Measurement: statistical modeling, observational analysis, metric design, forecasting, model evaluation
Data Platforms & Pipelines: GCP BigQuery, PostgreSQL, Hive, HDFS, Airflow, ETL, data pipelines
Communication: Tableau, dashboards, stakeholder-facing analysis for technical and non-technical audiences